On Confidence in Characterization and Application of Variation Models

Lerong Cheng, Puneet Gupta, and Lei He
University of California, Los Angeles
{lerong, puneet, lhe}@ee.ucla.edu

Abstract— **In this paper we study statistics of statistics. Statistical modeling and analysis have become the mainstay of modern design-manufacturing flows. Most analysis techniques assume that the statistical variation models are reliable. However, due to limited number of samples (especially in the case of lot-to-lot variation), calibrated models have low degree of confidence. The problem is further exacerbated when production volumes are low (≤ 65 lots) causing additional loss of confidence in the statistical analysis (since production only sees a small snapshot of the entire distribution). The problem of confidence in statistical analysis is going to be further worsened with advent of 450mm wafers. We mathematically derive the confidence intervals for commonly used statistical measures (mean, variance, percentile corner) and analysis (SPICE corner extraction, statistical timing). Our estimates are within 2% of simulated confidence values. Our experiments (with variability assumptions derived from test silicon data from a 45nm industrial process) indicate that for moderate characterization volumes (10 lots) and low-to-medium production volumes (15 lots), a significant guardband (e.g., 34.7% of standard deviation for single parameter corner, 38.7% of standard deviation for SPICE corner, and 52% of standard deviation for 95%-tile point of circuit delay) is needed to ensure 95% confidence in the results. The guardbands are non-negligible for all cases when either production or characterization volume is not large. We also study the interesting one production lot case which may be common for prototyping as well as for academic designs. The proposed methods require are not runtime-intensive (always within 10s) as they require Monte-Carlo simulations on closed form expressions.**

I. INTRODUCTION

With CMOS technology scaling down, process variation has become a major limiting factor for integrated circuit design. There are several works which analyze and model process variation (e.g., [1]–[5]). Variation models have been used in statistical leakage analysis (e.g., [6]–[13]) and statistical timing analysis (e.g., [14]–[24]) among others. Most statistical modeling and analysis assume that all the statistical characteristics of the variation sources, such as mean, variance, and skewness, are known and reliable. However, due to the limited measurement, the measured statistical characteristics may not be reliable. This is especially true for lot-to-lot variation, since the number of lots measured for characterization is usually small.

Moreover, the existing statistical analysis implicitly assume that the production chips have the same statistical characteristics as the corresponding population values. [1] When the number of production samples is not small, the production statistical characteristics may significantly deviate from their population values. Therefore, the uncertainty in statistics of measured data as well as production data should be considered in statistical analysis. [25] models the uncertainty of mean, variance, and correlation coefficients as an interval, and then estimates the range of mean and variance of circuit performance. However, this work only models the uncertainty as an interval, ignoring their true distributions. It also ignores the uncertainty introduced by limited number of production lots.

[1]Here by "population", we mean the statistical values in the case of infinite measured as well as production samples.

A. What is Confidence Interval

Before we move further, it is important to briefly review the concepts of confidence intervals in statistical analysis. Consider a standard normal random variable X, we choose n samples of it $(X_1, X_2, \ldots, X_n)$. Once the samples are chosen, the sample mean $\hat{\mu}$ and variance $\hat{\sigma}^2$ are fixed. However, if we repeatly choose n samples for 1,000 times, and calculate the sample mean and variance for each time, the results may differ. When the sample mean is smaller than a certain value μ_{conf} in 900 out of the 1000 runs, we say that we have 90% confidence that the sample mean is smaller than μ_{conf}. Notice that the confidence is not yield but the probability that certain statistical characteristic is within a given interval.. Once chips are produced, the distribution of (say) circuit delay is fixed. But due to the uncertainty of the statistical model, we do not know exactly the production mean and variance prior. For a given confidence, one can estimate the distribution of mean and variance (of course the actual mean and variance are going to be just one sample out of these distributions).

B. Contributions of This Paper

In this paper, we study the uncertainty of statistical models and analyze the impact of such uncertainty on statistical analysis. The contributions of this paper are:

1) Given the number of measured lots and the number of production lots,we estimate the distribution of the production mean and variance of variation sources.
2) For a confidence level, we estimate the worst case fast/slow corner ($\mu \pm k\sigma$ corner) for variation sources.
3) We extend the ideas to extract SPICE corners depending on desired confidence level as well as number of measured/produced silicon lots.
4) We estimate the confidence interval of mean, variance and quantile for statistical timing analysis.

Experimental results show that the required guardband (to reach desired confidence) estimated by our method is very close to the exact simulation value. We also observe that the guardband value increases dramatically when confidence level increases. We need to introduce up to 30% guardband value to achieve 95% confidence.

C. Practical Questions: Determining Confidence Induced Guardband for Real Designs

In this paper, we will propose techniques to quantify the confidence in statistical circuit analysis and estimate guardband required to attain a desired level of confidence. But it is intuitively that decreasing $\hat{n}$ and $\tilde{n}$ will reduce confidence in prediction made during design using statistical methods.

The answer to "what is the desired level of confidence" is not an easy one especially given the steep increase in guardband to ensure higher levels of confidence. Besides quality of models ($\hat{n}$) and production volume ($\tilde{n}$), two issues influence the choice of confidence level:

978-1-4244-5767-0/10/$26.00 2010 IEEE

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on August 14,2010 at 02:31:22 UTC from IEEE Xplore. Restrictions apply.

1) *Risk (in terms of schedule, performance, power, etc) tolerable for the design.* A lower confidence does *not* necessarily translate to lower yield but it implies that the probability of the statistical analysis actually matching real silicon is smaller. Therefore, confidence level is directly related to risk. For designs where power/performance constraints are flexible or time to volume is not critical (i.e., if yield is lower than expectation, running more lots is acceptable), a lower confidence is acceptable. Confidence levels provide a tradeoff between risk of higher manufacturing cost (due to low yield) and higher upfront design cost (due to guardbanding).
2) *Extent of post-silicon tuning options available.* Ability to change design characteristics during or post-manufacturing mitigates the need for high confidence in models. If the silicon foundry can extensively adjust the process (which essentially shifts the mean) for the design[2], the silicon can be *made* to match the analysis and ensuring high confidence in design-side estimates is not essential. Similarly, tunability using knobs such as adaptive body biasing, adaptive voltage scaling can also alleviate downside of low confidence.

In the end, the choice of confidence will be as much a business decision as a technical decision as it quantifies the desire for predictability of statistics of manufacturing.

D. Organization of This Paper

The rest of this paper is organized as follows: Section II gives the problem formulation; then Section III studies the confidence interval estimation for variation sources. Section IV presents the estimation of worst case delay for SPICE corner estimation and Section V analyzes the impact of mean and variance uncertainty on statistical timing analysis; and finally Section VI concludes this paper.

II. Problem Formulation

Process variation is decomposed into inter-lot (V_l), inter-wafer(V_w), inter-die (V_d), and within-die (V_r) variation:

$$\begin{aligned} V &= V_l + V_w + V_d + V_r \\ \mu_V &= \mu_l + \mu_w + \mu_d + \mu_r \\ \sigma_V^2 &= \sigma_l^2 + \sigma_w^2 + \sigma_d^2 + \sigma_r^2 \end{aligned} \quad (1)$$

where μ_l (σ_l^2), μ_w (σ_w^2), μ_d (σ_d^2), and μ_r (σ_r^2) are means (variances) of of inter-lot, inter-wafer, inter-die, and within-die variations, respectively. It has been shown that in case of finite number of samples the sample mean follows Gaussian distribution and the sample variance follow χ^2 distribution

$$\begin{aligned} \sigma_{\mu l}^2 &= \sigma_l^2/N_l & \sigma_{\sigma_l^2}^2 &= \sigma_l^2/N_l \\ \sigma_{\mu w}^2 &= \sigma_w^2/N_w & \sigma_{\sigma_w^2}^2 &= \sigma_w^2/N_w \\ \sigma_{\mu d}^2 &= \sigma_d^2/N_d & \sigma_{\sigma_d^2}^2 &= \sigma_d^2/N_d \\ \sigma_{\mu r}^2 &= \sigma_r^2/N_r & \sigma_{\sigma_r^2}^2 &= \sigma_r^2/N_r \end{aligned} \quad (2)$$

where N_l, N_w, N_d, and N_r are total number of lots, wafers, dies, and circuit elements. There are usually 20 to 50 wafers per lot, more than one hundred dies per wafer, and tens of measured circuit elements within each die. [3] Therefore, $N_r \gg N_d \gg N_w \gg N_l$. Hence $\sigma_{\mu l}^2 \gg \sigma_{\mu w}^2 \gg \sigma_{\mu d}^2 \gg \sigma_{\mu r}^2$, $\sigma_{\sigma_l^2}^2 \gg \sigma_{\sigma_w^2}^2 \gg \sigma_{\sigma_d^2}^2 \gg \sigma_{\sigma_r^2}^2$,. Therefore, the uncertainty of mean and variance comes largely from the lot-to-lot variation. In this paper, we focus our attention on lot-to-lot variation.

Table I illustrates five cases of number of measured lots and number of production lots.When the number of measured lots or production lots is small, the confidence interval is large and hence the statistical analysis are not reliable. In this case, we may want to guardband statistical analysis to ensure a certain degree of confidence. Example use models for these different scenarios can be: commodity process/high volume part (*L-L*); niche process/high column part (*S-L*); commodity process/niche part (*L-S*); niche process/niche part (*S-S*).

Case	# Measured lots	# Production lots	Confidence interval	Reliability of Statistical analysis
L-L	Large	Large	Small	High
S-L	Small	Large	Large	Low
L-S	Large	Small	Large	Low
S-S	Small	Small	Large	Low

TABLE I: Reliability of statistical analysis for different number of measured and production lots.

n	number of lots
μ, σ^2	mean and variance
$cn_{f/s}$	fast/slow corner
no hat	population value (μ, σ^2 and so on)
$\hat{\ }$	value obtain from measured lots ($\hat{n}$, $\hat{\mu}$, and so on)
$\tilde{\ }$	value of production lots ($\tilde{n}$, $\tilde{\mu}$, and so on)
$\square_t$	total value including inter-lot, inter-wafer, inter-die and within-die variation (μ_t, σ_t^2, and so on)
$\square_l$	inter-lot variation value (μ_l, σ_l^2, and so on)
$\square_w$	inter-wafer variation value (μ_w, σ_w^2, and so on)
$\square_d$	inter-die variation value (μ_d, σ_d^2, and so on)
$\square_r$	within-die variation value (μ_r, σ_r^2, and so on)
$\square_o$	total value except inter-lot variation (μ_o, σ_o^2, and so on) $\mu_o = \mu_w + \mu_d + \mu_r$, $\sigma_o^2 = \sigma_w^2 + \sigma_d^2 + \sigma_r^2$

TABLE II: Notations.

In practice, only the measured as opposed to the population value is known. Therefore, our the problem is: **Given the number of measured lots ($\hat{n}$) and production ($\tilde{n}$) lots and the measured values of mean ($\hat{\mu}$) and variance ($\hat{\sigma}^2$) of variation sources, estimate the distribution of statistical measures of production lots ($\tilde{\mu}$, $\tilde{\sigma}^2$).**

In rest of the paper, we assume that the lot-to-lot variation of all the variation sources follows Gaussian distribution.

III. Confidence Interval for Variation Sources

In this section, we will discuss confidence interval estimation for a single variation source. Table III summarizes the notation used in the rest of the paper. As should be clear from the discussion above, the measured valued is obtained deterministically from the measured chips. Our goal is to estimate the distribution of population value and production value from the given measured value.

A. Mean and Variance

Let's first discuss the mean and variance. From Equation (1), we calculate the total mean and variation of measured value as:

$$\hat{\mu}_t = \hat{\mu}_l + \mu_o \qquad \hat{\sigma}_t^2 = \hat{\sigma}_l^2 + \sigma_o^2 \quad (3)$$

As discussed in Section II, the uncertainty of mean and variance comes largely from lot-to-lot variation. Therefore, we assume the mean μ_o and variance σ_o^2 of all other variation are reliable. In the same way as above, we calculate the total mean and variance of the production as:

$$\tilde{\mu}_t = \tilde{\mu}_l + \mu_o \qquad \tilde{\sigma}_t^2 = \tilde{\sigma}_l^2 + \sigma_o^2 \quad (4)$$

In the rest of this subsection, we will focus on analyzing confidence interval for lot-to-lot variation.

In order to estimate the confidence interval, we first estimate the population values of mean and variance from the measurement data [26]:

$$\mu_l = \hat{\mu}_l + \hat{\sigma}_l \cdot M_1 \cdot \sqrt{\frac{\hat{n}-1}{\hat{n}Q_1}} \qquad \sigma_l^2 = (\hat{n}-1)\hat{\sigma}_l^2/Q_1 \quad (5)$$

where $M_1 \sim N(0,1)$ is a random variable with standard normal distribution, and $Q_1 \sim \chi^2_{\hat{n}-1}$ is a random variable with χ^2 distribution with $\hat{n}-1$ degrees of freedom [26]. Then, we estimate the

[2] This is usually possible only for high-volume, high-value designs.

[3] Note that 85 lots of 25 300mm wafers each with die-size of 100mm^2 amounts to production volume of 1.5 million chips.

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on August 14,2010 at 02:31:22 UTC from IEEE Xplore. Restrictions apply.

production mean and variance with respect to the population mean and variance:

$$\tilde{\mu}_l = \mu_l + \sigma_l \cdot M_2/\sqrt{\tilde{n}} \qquad \tilde{\sigma}_l^2 = \sigma_l^2 \cdot Q_2/(\tilde{n}-1) \tag{6}$$

where $M_2 \sim N(0,1)$ is a random variable with standard normal distribution and $Q_2 \sim \chi^2_{\tilde{n}-1}$ is a random variable with χ^2 distribution with $\tilde{n}-1$ degrees of freedom. Finally, we may combine Equations (5) and (6) to obtain the production mean ($\tilde{\mu}$) and variance ($\tilde{\sigma}^2$) from measured values ($\hat{\mu}$ and $\hat{sigma}^2$).

B. Simplifying the Large Lot Case

In the previous subsection, we consider the case that both $\hat{n}$ and $\tilde{n}$ are small (*S-S* case). We may simplify our model when either $\hat{n}$ or $\tilde{n}$ is large.

When the number of measured lots $\hat{n}$ is large (*L-S* case), we can approximate the population mean and variance as measured mean and variance:

$$\mu_l \approx \hat{\mu}_l \quad \sigma_l^2 \approx \hat{\sigma}_l^2$$

In this case, we may apply Equation (6) to estimate the distribution of production mean and variance.

In the other case when the number of production lots $\tilde{n}$ is large (*L-S* case), the production values of mean and variance are very close to the population value. I.e.,

$$\tilde{\mu}_l \approx \mu_l \quad \tilde{\sigma}_l^2 \approx \sigma_l^2$$

Therefore, we may only apply Equation (5) to estimate the distribution of production mean and variance.

C. Fast/Slow Corner Estimation

The next statistical characteristics we consider are fast and slow corners $\widetilde{cn}_{f/s}$ for variation sources.

Usually, the fast/slow corner is expressed as:

$$\widetilde{cn}_{f/s} = \tilde{\mu}_t \pm k_{f/s}\tilde{\sigma}_t \tag{7}$$

We may calculated the distribution of production mean $\tilde{\mu_t}$ and variance $\tilde{\sigma_t}$ from Equations (5) and (6), then calculate the distribution of fast/slow corner $\widetilde{cn}_{f/s}$.

For a given confidence level, we want to estimate the worst case fast/slow corner, $cn^w_{f/s}$. However, fast corner and slow corner are dependent on each other. Therefore, we need to handle them together. Assuming that hold time violation is harder to fix, higher confidence may be needed on fast corner. We estimate the worst case fast/slow corner such that there is cf_f confidence that the fast corner is larger than cn^w_f and there is cf_t confidence that the fast corner is larger than cn^w_f and the slow corner is smaller than cn^w_s, that is: [4]

$$P\{\widetilde{cn}_f > cn^w_f\} = cf_f \tag{8}$$
$$P\{\widetilde{cn}_f > cn^w_f, \widetilde{cn}_s < cn^w_s\} = cf_t$$

In this paper, we use Monte-Carlo simulations to obtain the joint distributions of the fast/slow corners. Then the worst case fast corner is calculated as:

$$cn^w_f = CDF^{-1}_{\widetilde{cn}_f}(cf_f) \tag{9}$$

With the worst case fast corner value, from the samples, we may also obtain the conditional CDF of $\widetilde{cn}_s$ given $\widetilde{cn}_f > cn^w_f$, $CDF_{\widetilde{cn}_s|\widetilde{cn}_f<cn^w_f}$. Then the worst case slow corner is calculated as:

$$cn^w_s = CDF^{-1}_{\widetilde{cn}_s|\widetilde{cn}_f<cn^w_f}(cf_t/cf_f) \tag{10}$$

We then express the worst case corner value as:

$$cn^w_{f/s} = \hat{\mu} \pm k'_{f/s}\hat{\sigma} \tag{11}$$

When the worst case corner values are known, we can easily obtain the value of $k'_{f/s}$.

[4] Without loss of generality, we assume fast implies larger parameter value (e.g. channel width).

cf_t %	cf_f %	k'_f	k'_s
50	60	3.080 (2.67%)	3.246 (8.20%)
60	70	3.155 (5.13%)	3.275 (9.13%)
70	80	3.255 (8.50%)	3.307 (10.2%)
80	90	3.422 (14.7%)	3.345 (11.5%)
90	95	3.594 (19.8%)	3.518 (17.2%)
95	99	4.042 (34.7%)	3.619 (20.6%)

TABLE III: Guardband of fast/slow corner for different degrees of confidence assuming $\hat{n} = 10$, $\tilde{n} = 15$.

cf_t %	cf_f %	k'_f	k'_s
50	60	3.035 (1.17%)	3.103 (3.43%)
60	70	3.112 (3.73%)	3.121 (4.03%)
70	80	3.124 (4.14%)	3.138 (4.60%)
80	90	3.236 (7.87%)	3.215 (7.17%)
90	95	3.401 (13.4%)	3.342 (11.4%)
95	99	3.625 (20.8%)	3.502 (16.7%)

TABLE IV: Guardband of fast/slow corner for different degrees of confidence assuming large $\hat{n}$ and $\tilde{n} = 15$.

Table III shows the value of $k'_{f/s}$ under different confidence levels. In the table, $100 \times (k'_{f/s} - k_{f/s})/k_{f/s}$ is indicative of the percentage guardband required to deal with limited data. In the experiment, we let $\hat{n} = 10$, $\tilde{n} = 15$, $k_f = k_s = 3$ and we assume that the variance of inter-lot variation is 18% of the total variance which is derived from wafer-to-wafer ring oscillator delay variation data from an industrial 45nm process with 348 wafers spread over 23 lots. We also assume that the variation source is with normal distribution [5]. In the experiment, we perform the estimation for 1000 runs. For each run, we generate $\hat{n}$ measured samples to obtain the measured mean $\hat{\mu}$ and variance $\hat{\sigma}^2$, and then apply the method above to calculate the worst case value of $k'_{f/s}$. Notice that the $k'_{f/s}$ depends on the measured values $\hat{\mu}$ and $\hat{\sigma}^2$, hence each run may result in a different $k'_{f/s}$ value. In the table, the $k'_{f/s}$ values are calculated as the average of 1000 runs. From the table, we can find that to ensure high confidence, $k'_{f/s}$ needs to be 30% over $k_{f/s}$, i.e., we need a large guardband.

As discussed in Section III-B, when the number of measured lots is large (*L-S*), the mean and variance can be calculated by only Equation (6). Hence the corner model can also be simplified. Table IV shows the average $k'_{f/s}$ value for the *L-S* case. As expected, the guardband in this case is smaller than the *S-S* case.

Similarly, when the number of production lots is large (*S-L*), we can also simplify the corner model by applying Equation (5) only. Table V shows the average $k'_{f/s}$ value for the *S-L* case. From the table, it is evident that the guardband is smaller than that of the *S-S* case.

Figure 1(a) compares the 90% confidence value of k'_f between the *L-S* (or *S-L*) and *L-L*. We observe that when $\tilde{n} \geq 60$ (or $\hat{n} \geq 85$), *L-S* (or *S-L*) can be looked on as *L-L*. That is, the statistical analysis is reliable and we do not need to perform guardband estimation. In the experiment, when error of k'_f value between *L-L* and *L-S* (or *S-L*) case are within 3%, we consider $\hat{n}$ (or $\tilde{n}$) value is large.

Note that increasing lot-to-lot variation will increase the value of $\hat{n}.\tilde{n}$ that can be considered large as shown in Figure 1(b). This is

[5] The nominal mean and variance of variation sources does not affect the value of k'_f/s

cf_t %	cf_f %	k'_f	k'_s
50	60	3.052 (1.73%)	3.132 (4.40%)
60	70	3.134 (4.47%)	3.155 (5.17%)
70	80	3.165 (5.50%)	3.193 (6.43%)
80	90	3.272 (9.07%)	3.267 (8.90%)
90	95	3.431 (14.4%)	3.370 (12.3%)
95	99	3.690 (23.0%)	3.523 (17.4%)

TABLE V: Guardband of fast/slow corner for different confidence levels assuming $\hat{n} = 10$ and large $\tilde{n}$.

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on August 14,2010 at 02:31:22 UTC from IEEE Xplore. Restrictions apply.

because when the lot-to-lot variation increase, the uncertainty of mean and variance increases. Therefore, we need a larger number of measured lots (or production lots) to be considered as large.

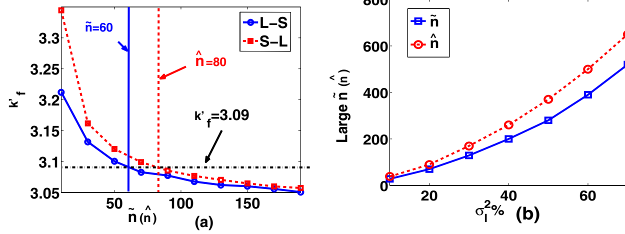

Fig. 1: (a) Comparison of *S-L*, *L-S*, and *L-L* case. (b) Number of $\hat{n}$ or $\tilde{n}$ which can be considered as large for different fraction of lot-to-lot variation assuming maximum 3% error at 90% confidence.

IV. SPICE FAST/SLOW CORNER

Fast/slow corners for circuit simulation is the major interface between design and process. A common approach to derive the corners is to use measured values from canonical circuits such as an inverter chain. Variation source values corresponding to the corner delay $D^w_{f/s}$ are then calculated.

We assume that the inverter chain delay is a linear function of variation sources:

$$D = d_0 + \sum_{i=1}^{m} c_i X_i \tag{12}$$

where m is the number of variation sources, d_0 is the nominal delay, X_i's are variation sources, and c_i's are sensitivity coefficients of inverter chain delay to variation sources. Considering all variation sources to be independent, the mean and variance of the product inverter chain delay can be calculated as:

$$\begin{aligned} \tilde{\mu}_D &= d_0 + \sum_{i=1}^{m} c_i(\mu_{oi} + \tilde{\mu}_{li}) \\ \tilde{\sigma}^2_D &= \sum_{i=1}^{m} c_i^2(\sigma^2_{oi} + \tilde{\sigma}^2_{li}) \end{aligned} \tag{13}$$

where μ_{oi} and σ^2_{oi} are the mean and variance of other variation for the ith variation source, respectively, $\tilde{\mu}_{li}$ and $\tilde{\sigma}^2_{li}$ are the mean and variance of lot-to-lot variation for the ith variation source, respectively. $\tilde{\mu}_{li}$ and $\tilde{\sigma}^2_{li}$ can be calculated as shown in Section III-A.

Usually, the delay corner is expressed as

$$\tilde{D}_{f/s} = \tilde{\mu}_D \pm k_{f/s}\tilde{\sigma}_D \tag{14}$$

As in Section III-C we use Monte-Carlo simulation to obtain the PDF of the fast and slow delay corners. One sided confidence interval is then used to obtain the worst case fast and slow delay corners D^w_f/s. Corresponding variation sources corner values are calculated by solving the following [27]:

$$\begin{aligned} D^w_{f/s} &= d_0 + \sum_{i=1}^{m} c_i X^w_i \\ \frac{X^w_{1f/s} - \hat{\mu}_1}{c_1\hat{\sigma}_1} &= \frac{X^w_{2f/s} - \hat{\mu}_2}{c_2\hat{\sigma}_2} \ldots = \frac{X^w_{mf/s} - \hat{\mu}_m}{c_m\hat{\sigma}_m} \end{aligned} \tag{15}$$

Table VI illustrates the guardband percentage for the worst case delay and the corresponding variation source values under different confidence levels. In the table, the guardband percentage is calculated as: 100×|worst_case_value - nominal_value|/nominal_value. In the experiment, we use PTM bulk low power SPICE model for 45nm technology [28]. We assume three variation sources, gate length L, threshold voltage for NMOS V_{tn} and PMOS V_{tp}. For gate length variation, we assume $3\sigma = 3nm$, for threshold voltage variation, we assume that the 3σ value is 20% of the nominal value. We also assume that the inter-lot variation is 18% as in previous sections. In the experiment, we let $\hat{n} = 10$ and $\tilde{n} = 15$. Since measured corner values affect production corners,we show the average of 1000 runs.

From the table, we observe that the under the same confidence level, the guardband percentage of delay corner is less than that of

cf_t	cf_f	D^w_f	L^w_f	V^w_{tnf}	V^w_{tpf}	D^w_s	L^w_s	V^w_{tns}	V^w_{tps}
50	60	0.29	0 07	0.21	0.20	0.52	0.13	0.39	0.36
60	70	0.57	0 14	0.43	0.40	0.77	0.19	0.58	0.55
70	80	1.15	0 29	0.86	0.81	1.29	0.32	0.97	0.91
80	90	2.01	0 50	1.50	1.41	1.80	0.45	1.35	1.27
90	95	3.15	0 79	2.36	2.22	2.84	0.71	2.13	2.00
95	99	4.58	1 15	3.44	3.23	3.61	0.90	2.71	2.54
Nominal		349	42.62	0.558	0.603	388	47.41	0.612	0.643

TABLE VI: Percentage guardband of worst case delay corner and corresponding variation source corners under different confidence levels assuming $\hat{n} = 10$, $\tilde{n} = 15$. Note: delay value is in ps, L_{gate} value is in nm, and V_{th} value is in V.

k'_f value of variation sources as shown in Table III. This is because the k'_f value of variation sources does not depend on the nominal mean and variance. However, for the worst case delay, the guard band percentage does depend on nominal mean. When the nominal mean increases, the guardband percentage decreases.

In Table VII, we also compare the targeted confidence and the exact confidence of the estimated worst case corner value. The flow to obtain the simulated confidence is shown in Figure 2. In the experiment, we let $n_{run} = 1,000$ From the table, we find that the exact simulated confidence is a little bit higher than the target value. Such error largely comes from the fitting of the linear delay model to SPICE.

1. $n_f = 0$, $n_t = 0$
2. for i=1 to n_{run}
   3. Generate $\hat{n}$ samples of measured lots /* calculate estimated value */
   4. Calculate $\hat{\mu}$ and $\hat{\sigma}^2$ for each variation sources from samples
   5. Perform 10,000-sample MC simulation on Equation (14) according to $\hat{\mu}$ and $\hat{\sigma}^2$.
   6. Obtain worst case corners: D^w_f, D^w_s.
   7. Generate $\tilde{n}$ samples of production lots /* calculate simulated value */
   8. Calculate $\tilde{\mu}$ and $\tilde{\sigma}^2$ for each variation source from samples.
   9. Perform 10,000-sample SPICE MC simulation on inverter chain delay based on $\tilde{\mu}$ and $\tilde{\sigma}^2$.
   10. Calculate mean μ_D and variance σ^2_D of SPICE MC samples.
   11. Calculate simulated corners: $D_f = \mu_D - k_f\sigma_D$, $D_s = \mu_D + k_s\sigma_D$.
   12. if $D_f > D^w_f$ /* compare simulated value and estimated value */
      13. n_f=$n_f + 1$.
      14. if $D_s < D^w_s$
         15. $n_t = n_t + 1$
16. $cf_f = n_f/n_{run}$, $cf_t = n_t/n_{run}$

Fig. 2: Flow to simulate confidence.

Targeted $Conf$		Simulated $Conf$	
cf_t	cf_f	cf_t	cf_f
50	60	51.9	61.6
60	70	61.7	72.4
70	80	71.6	81.4
80	90	81.1	91.3
90	95	90.8	95.7
95	99	95.5	99.2

TABLE VII: Confidence comparison for inverter chain based corner assuming $\hat{n} = 10$, $\tilde{n} = 15$.

Table VIII and IX show the average guardband percentage for the *L-S* and *S-L* cases, respectively. Note that the guardband, though appears to be small, is significant when compared to variation itself (e.g. σ for V_{tp} variation is only 6.7% of nominal value, while the guardband of fast corner is up to 2.58% of nominal value. That is, the guardband value is up to 38.7% of standard deviation).

Figure 1 compares the 90% confidence value of D^w_f between the *L-S* (or *S-L*) and *L-L*. In the figure, D^w_f is normalized with respect to the nominal value. From the figure, We observe that $\tilde{n} \geq 30$ (or $\hat{n} \geq 45$), can be considered as "large". Similar to the single source case, when error of k'_f value between *L-L* and *L-S* (or *S-L*) case are within 3%, we consider $\hat{n}$ (or $\tilde{n}$) value is large. We also find that for the worst case delay corner, the value of $\hat{n}$ (or $\tilde{n}$) to be considered as

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on August 14,2010 at 02:31:22 UTC from IEEE Xplore. Restrictions apply.

cf_t	cf_f	D_f^w	L_f^w	V_{tnf}^w	V_{tpf}^w	D_s^w	L_s^w	V_{tns}^w	V_{tps}^w
50	60	0.21	0.05	0.16	0.15	0.39	0.10	0.29	0.27
60	70	0.43	0.11	0.32	0.30	0.58	0.14	0.43	0.41
70	80	0.86	0.21	0.64	0.61	0.97	0.24	0.72	0.68
80	90	1.50	0.38	1.13	1.06	1.35	0.34	1.01	0.95
90	95	2.36	0.59	1.77	1.67	2.13	0.53	1.59	1.50
95	99	3.44	0.86	2.58	2.42	2.71	0.68	2.03	1.91

TABLE VIII: Percentage guardband of worst case delay corner and corresponding variation sources corners under different confidence levels assuming large and $\tilde{n} = 15$

cf_t	cf_f	D_f^w	L_f^w	V_{tnf}^w	V_{tpf}^w	D_s^w	L_s^w	V_{tns}^w	V_{tps}^w
50	60	0.24	0.06	0.18	0.17	0.44	0.11	0.33	0.31
60	70	0.49	0.12	0.37	0.34	0.66	0.16	0.49	0.46
70	80	0.97	0.24	0.73	0.69	1.10	0.27	0.82	0.77
80	90	1.70	0.43	1.28	1.20	1.53	0.38	1.15	1.08
90	95	2.68	0.67	2.01	1.89	2.41	0.60	1.81	1.70
95	99	3.90	0.97	2.92	2.75	3.07	0.77	2.30	2.16

TABLE IX: Percentage guardband of worst case delay corner and corresponding variation source corners under different confidence levels assuming $\hat{n} = 10$ and large $\tilde{n}$

large is smaller than that of the single variation source as in Figure 1. This is because the guardband of worst case delay is smaller than that of single variation source as discussed above. Hence, we need fewer lots to achieve the same model reliability.

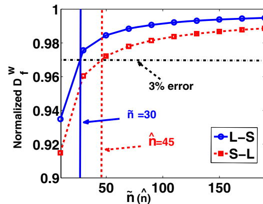

Fig. 3: 90% confident D_f^w with varying $\hat{n}$ and $\tilde{n}$.

In this section, we illustrated the extra guardband required in corners (which ironically are themselves guardbands by definition) to ensure confidence in statistical analysis when the number of lots used to characterize variation or number of production silicon lots are small. Next we discuss the impact of limited characterization or production data on one of the most talked about chip-level statistical analysis methods, namely, statistical timing analysis.

V. Confidence in Statistical Timing Analysis

In this section, we will discuss about confidence estimation for statistical static timing analysis (SSTA). We assume that all variation sources normally distributed, and apply the linear SSTA method [29]. By performing SSTA, the circuit output delay is expressed as a linear canonical form:

$$D = d_0 + \sum_{i=1}^{m} c_i X_i \tag{16}$$

Since circuit delay has a linear canonical form, its mean and variance can be calculated in the same way as shown in Equation (13). We again use Monte-Carlo simulation on the linear equation to obtain the CDF of the output mean ($CDF_{\mu_D}(\cdot)$) and variance ($CDF_{\sigma_D^2(\cdot)}$). Then for a given confidence level $Conf$, it is easy to obtain the worst case mean and variance: [6]

$$\begin{aligned} \mu^w &= CDF_{\mu_D}^{-1}(Conf) \\ \sigma^{w2} &= CDF_{\sigma_D^2}^{-1}(Conf) \qquad \sigma^w = \sqrt{\sigma^{w2}} \end{aligned} \tag{17}$$

Another quantity of interest is a certain percentile point. Since we assume Gaussian variation sources and apply linear canonical form to approximate circuit delay, the circuit delay also follows Gaussian distribution. Therefore, the percentile point $C(p\%)$ can be calculated as:

$$\begin{aligned} C(p\%) &= \mu + k\sigma \\ k &= \Phi^{-1}(p\%) \end{aligned} \tag{18}$$

where $\Phi(\cdot)$ is the CDF of standard normal distribution. In this way, the percentile point is converted to a $\mu + k\sigma$ corner. We may apply the same method as in Section IV to obtain the CDF of the percentile point, and then calculate its worst case value.

[6] Notice that we have known the distribution of mean and variance, it is easy for us to obtain any interval of mean and variance under a certain confidence.

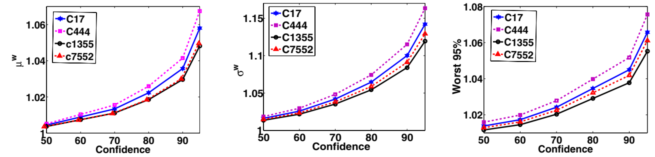

Fig. 4: Worst case mean, standard deviation, and 95% percentile point normalized with respect to the nominal value for ISCAS85 benchmarks, assuming $\hat{n} = 10$ and $\tilde{n} = 15$.

Targeted $conf$	Simulated $conf$ μ	σ^2	95%
50	51.3	51.4	51.3
60	60.8	60.9	61.2
70	70.1	70.7	70.8
80	80.9	81.0	80.3
90	90.5	90.4	90.2
95	95.5	95.2	95.3

TABLE X: Confidence comparison for ISCAS85 benchmark 95% percentile point delay assuming $\hat{n} = 10$, $\tilde{n} = 15$.

Figure 4 illustrates the worst case mean, standard deviation, and 95% percentile point for ISCAS85 benchmarks. In the experiment, we apply the same experimental setting as that for the inverter chain as in Section IV. As discussed before, the worst case value depends on the measured values of mean and variance. Hence, different runs may result in different worst case values. In the figure, the worst case value is averaged over 1,000 runs and normalized with respect to the nominal value. From the figure, we observe that the guardband of chip delay is lower than that of single variation source and inverter chain delay. This is because chip delay has a large nominal mean to variance ratio than inverter chain. However, although the guardband value seems not large, it is very significant compared to the scale of variation. For example, for the 95% percentile point, the nominal value is 6.9X of standard deviation and the guardband value is up to 7.6% of nominal value to ensure 95% confidence. That means, the guardband value is up to 52% of standard deviation.

In Table X, we compare the targeted confidence and the simulated confidence of the estimated confidence interval. The simulated confidence is obtained in Figure 2. The only difference is that instead of running SPICE Monte-Carlo simulation, we apply SSTA [24] to obtain the chip delay distribution. Due to our simplifying assumption that the linear canonical form is independent of small changes in mean and variance, there is a small error between targeted and simulated confidence.

Figure 5 illustrate the mean, standard deviation, and 95% percnetile point of ISCAS85 benchmarks under different confidence level for *S-L* case. We ignore *L-S* case where the need of SSTA is not well motivated.

Figure 6 compares the 90% confidence value of 95% percentile point for C7552 between *S-L* and *L-L*. In the figure, percentile point is normalized with respect to the nominal value. From the figure, we observe that the guardband of chip delay is lower than that of single variation source and inverter chain delay. We find that we only need a small number of lots ($\hat{n} \geq 24$) to bound the error between *S-L* and

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on August 14,2010 at 02:31:22 UTC from IEEE Xplore. Restrictions apply.

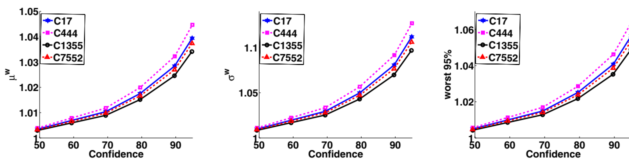

Fig. 5: Normalized worst case mean, standard deviation, and 95% percentile point for ISCAS85 benchmarks, assuming $\hat{n} = 10$ and $\tilde{n}$ is large.

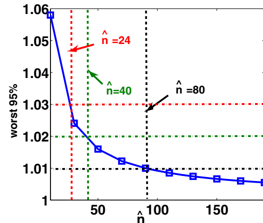

Fig. 6: 90% confidence worst case 95% percentile point with varying $\hat{n}$ for C7552. Percentile point is normalized with respect to the nominal value.

L-L under 3%. If we want to bound the error to 2% or 1%, we need $\hat{n} \geq 40$ and $\hat{n} \geq 80$, respectively.

Again, we see that limited data can lead to significantly large guardbanding to ensure good confidence in analysis results. The method discussed in this section is fairly straightforward, fast (essentially Monte Carlo on a linear expression, runtime is within 10s) and accurate to estimate confidence in results of statistical timing analysis especially when the variability models are known to suffer from limited lot-to-lot variation data.

VI. Conclusions and Ongoing Work

In this paper, we have studied impact of characterization and production silicon volumes on believability of statistical variation models and analysis. We have developed methods to estimate the distribution of production silicon statistical characteristics (e.g., mean, variance, percentile corners, etc) in terms of (unreliable) measured statistical characteristics, measured data volume and intended production volume. The loss of reliability largely stems from lot-to-lot variation. Our experiments and results indicate the following.

1) For small characterization or production volumes, best/worst corners for a variation source may need as much as 30 % guardband to attain 95% confidence. Our estimates show strong agreement with confidence measured from a set of real silicon data (maximum error $\leq$ 4%).
2) For a typical SPICE corner extraction methodology, the guardband needed for 95% confidence is about 14% of the nominal uncertainty (difference between best/worst corners) for low-to-medium volume production ($\hat{n} = 10, \tilde{n} = 15$).
3) Finally, the 95th percentile delay as estimated using statistical timing analysis needs a guardband of 5%-7%.

Predicted statistics are reliable for measured lots are larger than 85 and number of production lots larger than 60. We assumed lot-to-lot variation to be 18% of total variance for our experiments. The required guardband will increase if lot-to-lot variation increases as a fraction of total variation. We believe that confidence-level induced guardband should be considered for all low-to-medium volume designs. Moreover, the problem is likely to become more severe when 450mm wafer sizes are adopted by the industry (since less number of lots would be required to meet the same production volume).

Our ongoing work is trying to extend the work to statistical power calculations. Further, we plan to work out simpler approximations to estimate the guardband without Monte-Carlo simulations.

References

[1] J. Xiong, V. Zolotov, and L. He, “Robust extraction of spatial correlation,” in *ISPD*, Apr 2006.
[2] F. Liu, “A general framework for spatial correlation modeling in vlsi design,” in *DAC*, Jun 2007.
[3] J.-H. Liu, M.-F. Tsai, L. Chen, and C. C.-P. Chen, “Accurate and analytical statistical spatial correlation modeling for vlsi dfm applications,” in *DAC*, Jun 2008.
[4] T. Sato, H. Ueyama, N. Nakayama, and K. Masu, “Determination of optimal polynomial regression function to decompose on-die systematic and random variations,” in *ASPDAC*, Feb 2008.
[5] S. Reda and S. R. Nassif, “Analyzing the impact of process variations on parametric measurements: Novel models and applications,” in *DATE*, 2009.
[6] Y. Lu and V. Agrawal, “Statistical leakage and timing optimization for submicron process variation,” in *VLSI Design*, Jan 2007.
[7] S. Bhardwaj and S. Vrudhula, “Leakage minimization of digital circuits using gate sizing in the presence of process variations,” *TCAD*, Mar 2008.
[8] R. Chen and H. Zhou, “Fast estimation of timing yield bounds for process variations,” *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, Mar 2008.
[9] X. Ye, Y. Zhan, and P. Li, “Statistical leakage power minimization using fast equi-slack shell based optimization,” in *DAC*, 2007.
[10] X. Shen, Y. Wang, R. Luo, and H. Yang, “Leakage power reduction through dual vth assignment considering threshold voltage variation,” in *ASICON*, Oct 2007.
[11] L. Tao and Y. Zhiping, “Statistical analysis of full-chip leakage power considering junction tunneling leakage,” in *DAC*, Jun 2007.
[12] X. Li, J. Le, and L. Pileggi, “Projection based statistical analysis of full-chip leakage power with non-log-normal distributions,” in *DAC*, Jun 2006.
[13] D. H.F., S.-C. Lin, and B. K, “A statistical framework for estimation of full-chip leakage-power distribution under parameter variations,” *Electron Devices, IEEE Transactions on*, Nov 2007.
[14] A. Ramalingam, A. K. Singh, S. R. Nassif, G.-J. Nam, M. Orshansky, and D. Z. Pan, “An accurate sparse matrix based framework for statistical static timing analysis,” in *ICCAD*, Nov 2006.
[15] S. Abbaspour, H. Fatemi, and M. Pedram, “Parameterized block-based non-gaussian statistical gate timing analysis,” in *ASPDAC*, Jan 2006.
[16] L. Zhang, Y. Hu, and C.-P. Chen, “Statistical timing analysis with path reconvergence and spatial correlations,” in *DATE*, Mar 2006.
[17] S. Abbaspour, H. Fatemi, and M. Pedram, “Non-gaussian statistical interconnect timing analysis,” in *DATE*, Mar 2006.
[18] F. Gong, W. Yu, Z. W. Z. Yu, and C. Yan, “Efficient techniques for 3-d impedance extraction using mixed boundary element method,” in *ASPDAC*, 2008.
[19] J. Xiong, Y. Shi, V. Zolotov, and C. Visweswariah, “Multilayer process space coverage for at-speed test,” in *DAC*, 2009.
[20] J. Xiong, Y. Shi, V. Zolotov, and C. Visweswariah, “Pre-atpg path selection for near optimal post-atpg process space coverage,” in *ICCAD*, 2009.
[21] C. Liu, J. Su, and Y. Shi, “Temperature aware routing synthesis considering spatiotemporal hotspot,” in *ICCD*, 2008.
[22] Y. Zhan, A. J. Strojwas, D. Newmark, and M. Sharma, “Generic statistical timing analysis with non-gaussian process parameters,” in *Austin Conference on Integrated Systems and Circuits*, May 2006.
[23] L. Cheng, J. Xiong, and L. He, “Non-linear statistical static timing analysis for non-gaussian variation sources,” in *DAC*, Jun 2007.
[24] L. Cheng, J. Xiong, and L. He, “Non-gaussian statistical timing analysis using Second-Order polynomial fitting,” in *ASPDAC*, Feb 2008.
[25] G. Yu, W. Dong, Z. Fang, and P. Li, “Statistical static timing analysis considering process variation model uncertainty,” *TCAD*, Oct. 2008.
[26] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1965.
[27] M. Orshansky, S. R. Nassif, and D. Boning, *Design for Manufacturability and Statistical Design*. Springer, 2007.
[28] “PTM SPICE model,” in *http://www.eas.asu.edu/ ptm/latest.html*.
[29] C. Visweswariah, K. Ravindran, K. Kalafala, S. Walker, and S. Narayan, “First-order incremental block-based statistical timing analysis,” in *DAC 04*, June 2004.

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on August 14,2010 at 02:31:22 UTC from IEEE Xplore. Restrictions apply.